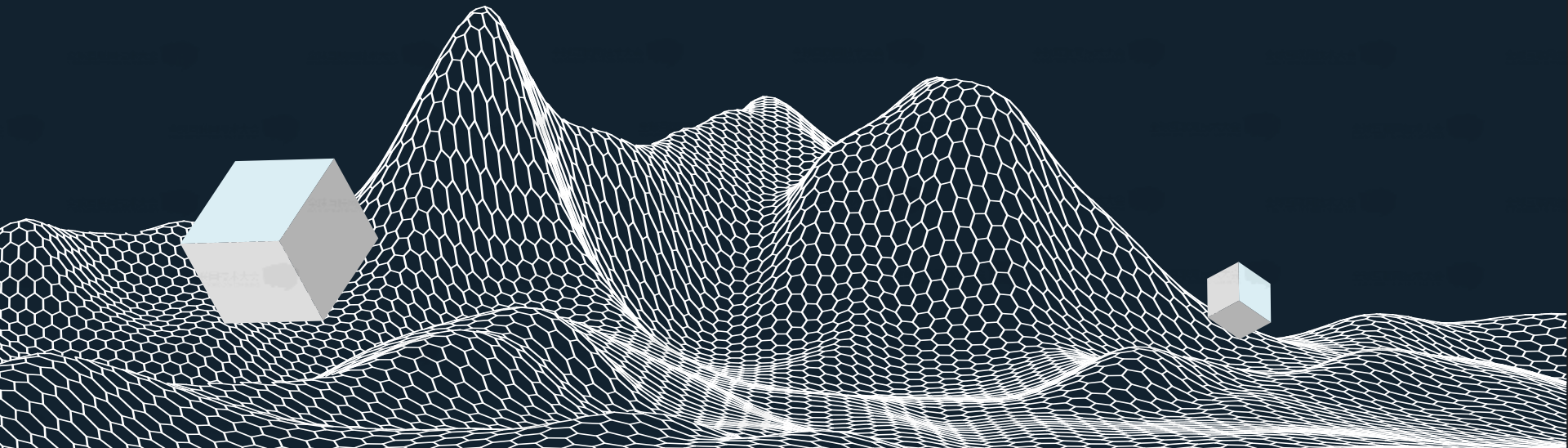


# 利用大数据治理激活数据

丛磊 白山合伙人兼工程副总裁



1

大数据的生与死

2

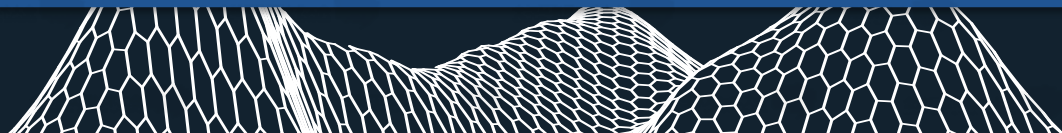
数据流动起来的技术栈

3

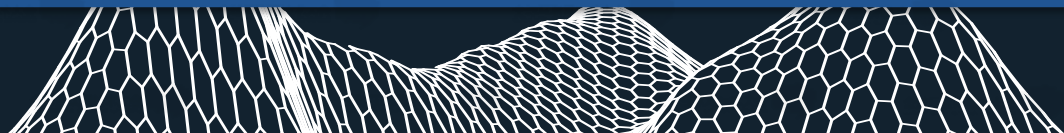
数据开放的核心流程

4

云聚合产品功能列表

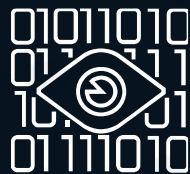


# 1 大数据的生与死

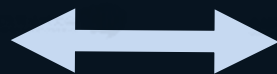


# 大数据的“生”与“死”

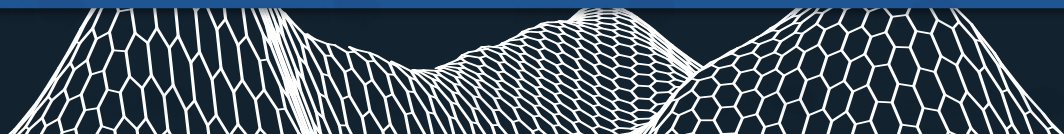
- 流动的数据产生价值
- 99%的大数据并没有流动起来
- 数据为什么无法流动



数据



用户



# 数据无法流动的原因

## 没有数据接口

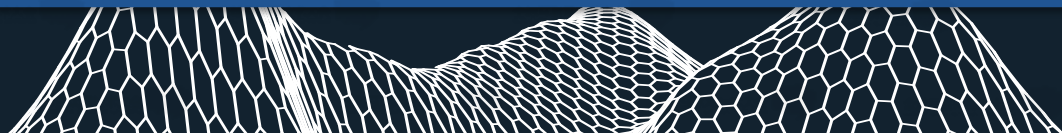
- 传统数据文件没有接口
- 数据库没有接口

## 接口不具备对外访问能力

- 接口不规范
- 缺乏认证体系
- 访问速度低下
- 缺少数据统计分析

## 接口很难调用

- 缺乏开发文档
- 缺少SDK
- 很难调试



# 如何让数据“活”起来

## ➤ 重新激活你的数据

- 数据开放
- 信息共享

## ➤ 政府、企业的OpenAPI开放平台

- 美国政府数据开放与交换中心 data.gov
- BBC开放平台
- Coca-cola开放平台

国务院关于印发政务信息资源共享管理暂行办法的通知

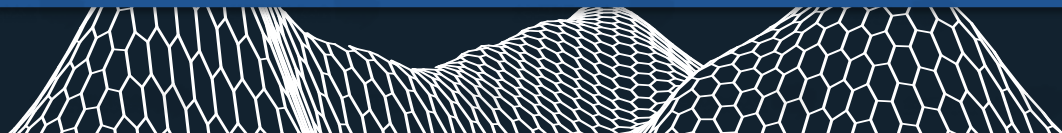
国发〔2016〕51号

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：

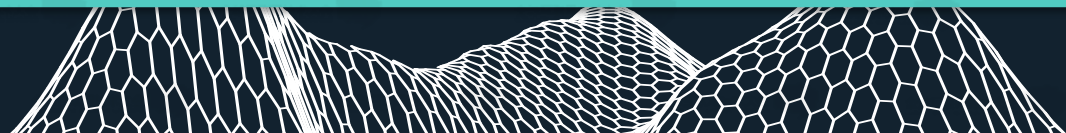
现将《政务信息资源共享管理暂行办法》印发给你们，请认真贯彻执行。

国务院

2016年9月5日



## 2 数据流动起来的技术栈



技术需求

便捷

安全

快速

通用

可靠

OpenAPI 平台

移动APP

IOT多终端

政府数据开放与共享整体技术架构

核心技术栈

传输加速

分发加速

事件驱动加速

DNS加速

数据安全

权限控制

网络隔离

数据加密

数字签名

接口规范

XML

JSON

SOAP

网络协议

HTTP

TCP

UDP

数据存储

数据库

文件系统

分布式存储

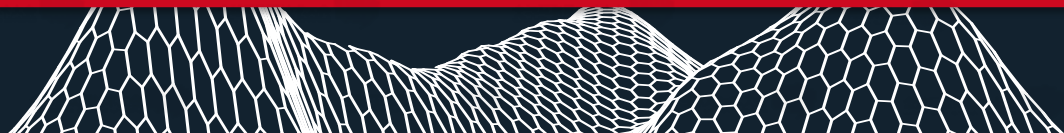
大数据分析

机器学习

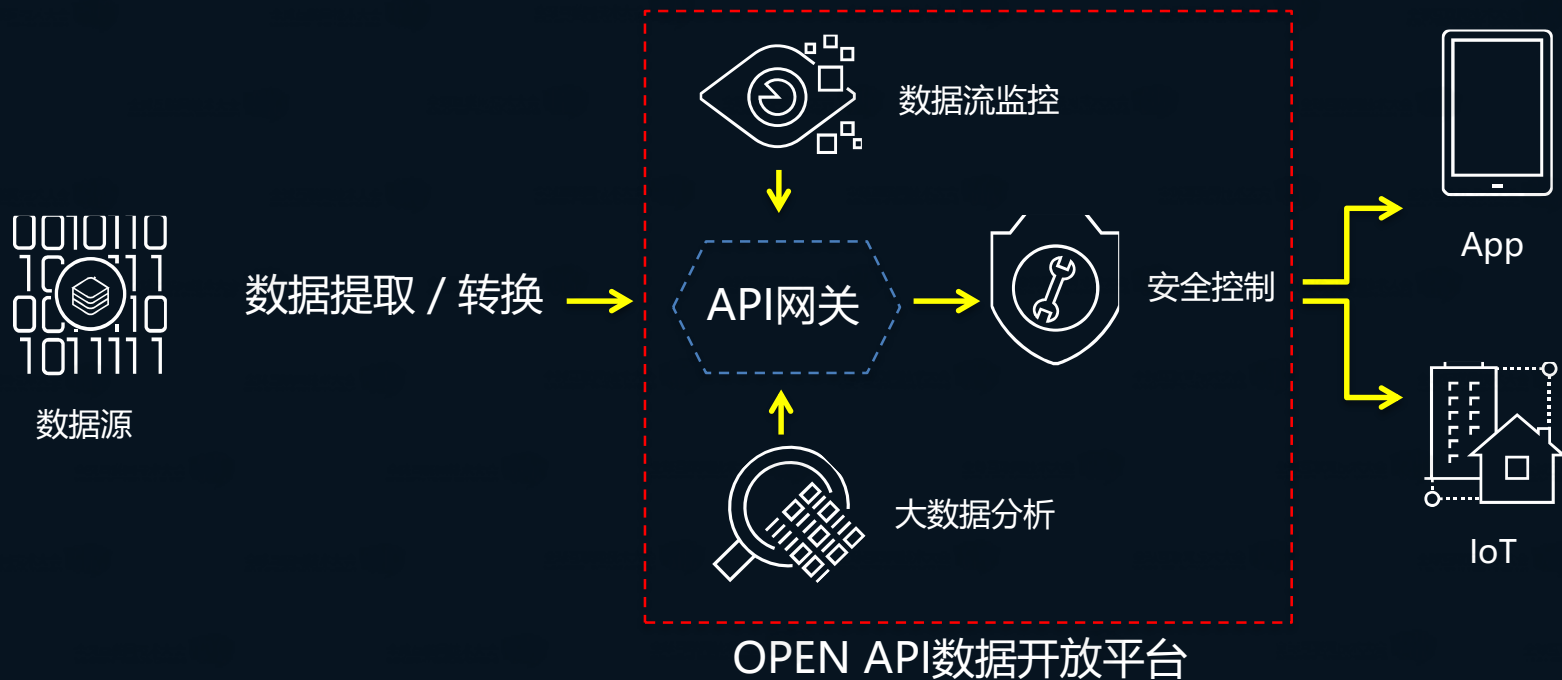
实时计算



# 3 数据开放的核心流程

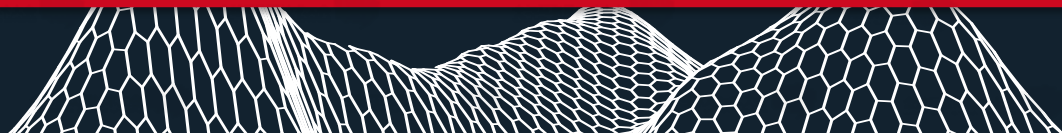


# 数据开放的核心流程



# 数据开放的核心技术

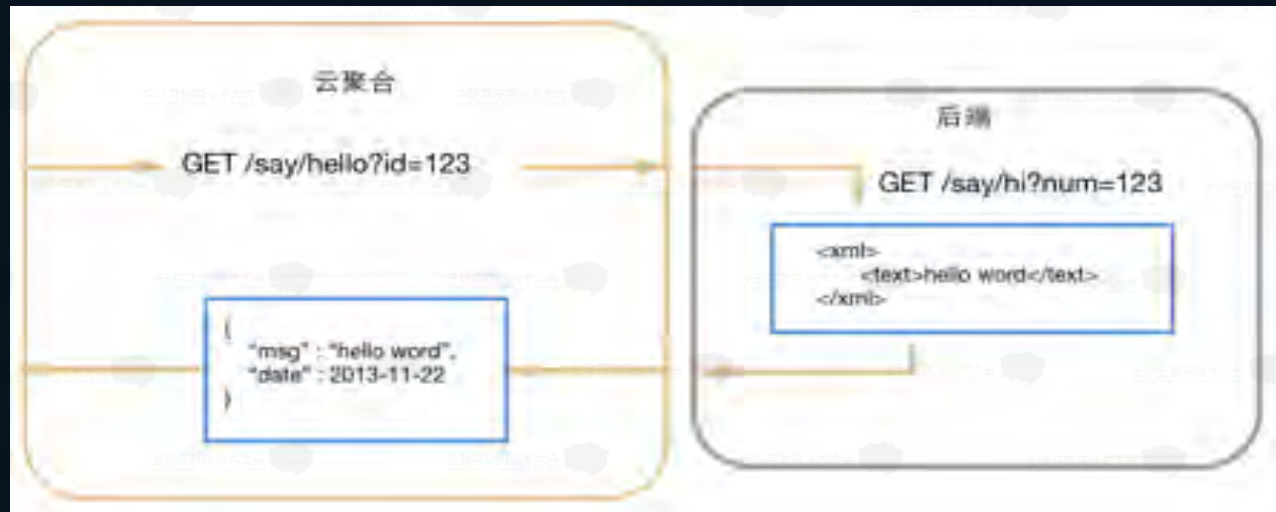
- API接口化
  - RESTful
  - Swagger
- 构造移动端的API
  - 精准请求
  - 响应式设计
- 如何应对不标准的API？



# 数据开放的核心技术

## ➤ API适配

- Query String
- Path
- Header
- Body



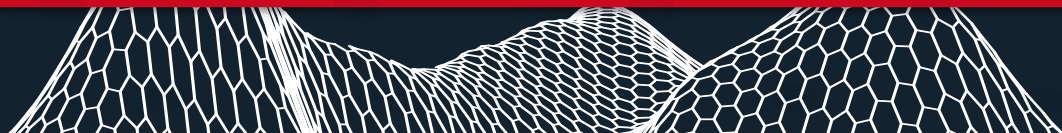
# 数据开放的核心技术

## ➤ API自动生成

- Database
- FILE
- HTML



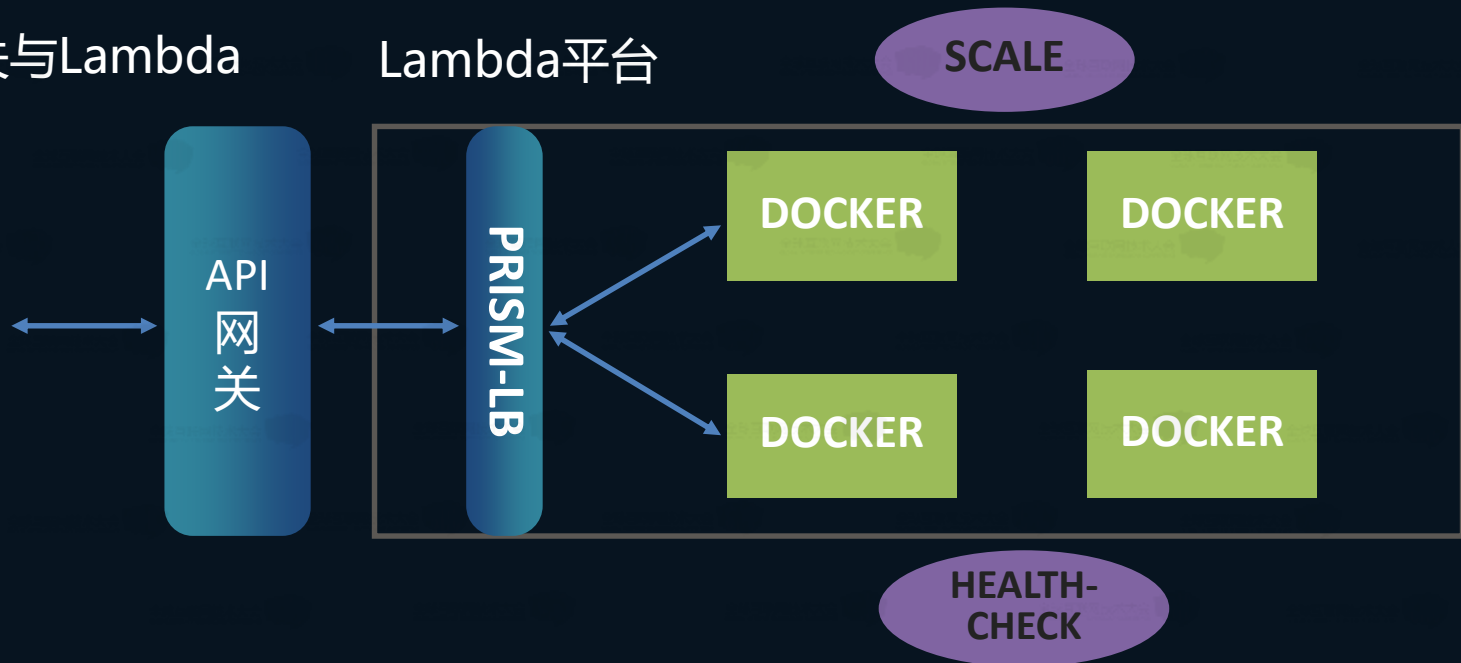
## ➤ ETL ( Extract Transform Loading ) 到 ETI ( Extract Transform Interface )



# 数据开放的核心技术

➤ API网关与Lambda


Lambda平台




# 数据开放的核心技术

## ➤ API网关与Lambda

### 场景A



智能硬件向云端POST数据，API网关通过Lambda进行数据筛选、聚合，然后向云端发送，大大降低数据传输量。



用户向云端GET数据，API网关将访问日志推送到大数据分析平台进行分析，并将分析结果汇总发送邮件。

### 场景B

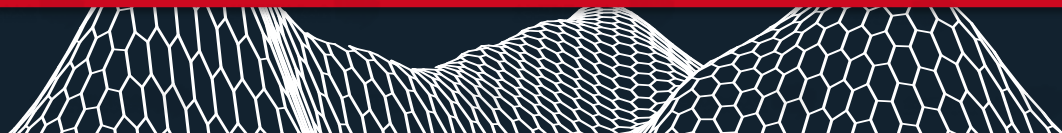
# 数据开放的核心技术

## ➤ API防攻击

## DDOS攻击 VS CC攻击

- CC攻击
- 特征识别
- 深度学习

	来源	危害	协议	变种
CC攻击	真实IP	业务故障、或者业务可用 但用户隐私等被窃取	HTTP	较多
DDoS攻击	伪造IP	流量打满，业务彻底不可用	TCP/UDP/IP	较少

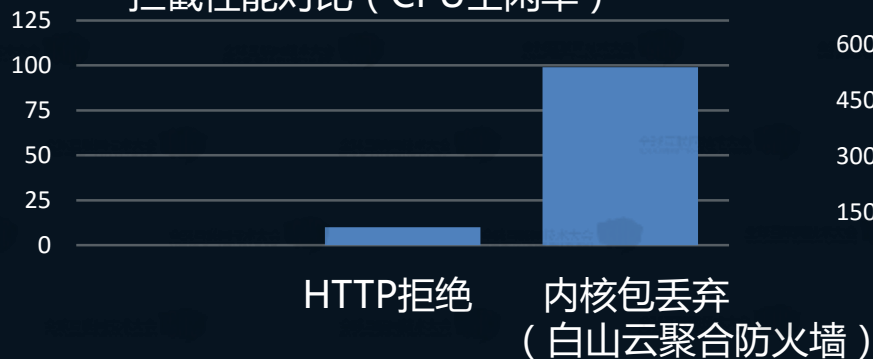




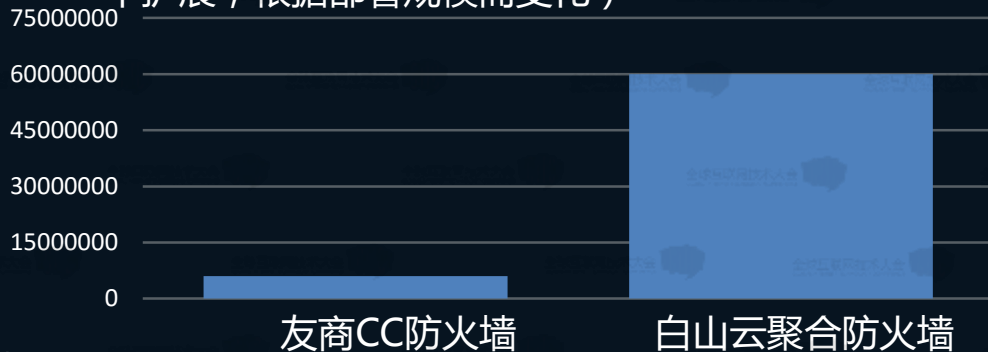
# 数据开放的核心技术

## ➤ 性能指标

拦截性能对比 (CPU空闲率)



分析性能对比 (请求数/分钟, 性能支持水平扩展, 根据部署规模而变化)



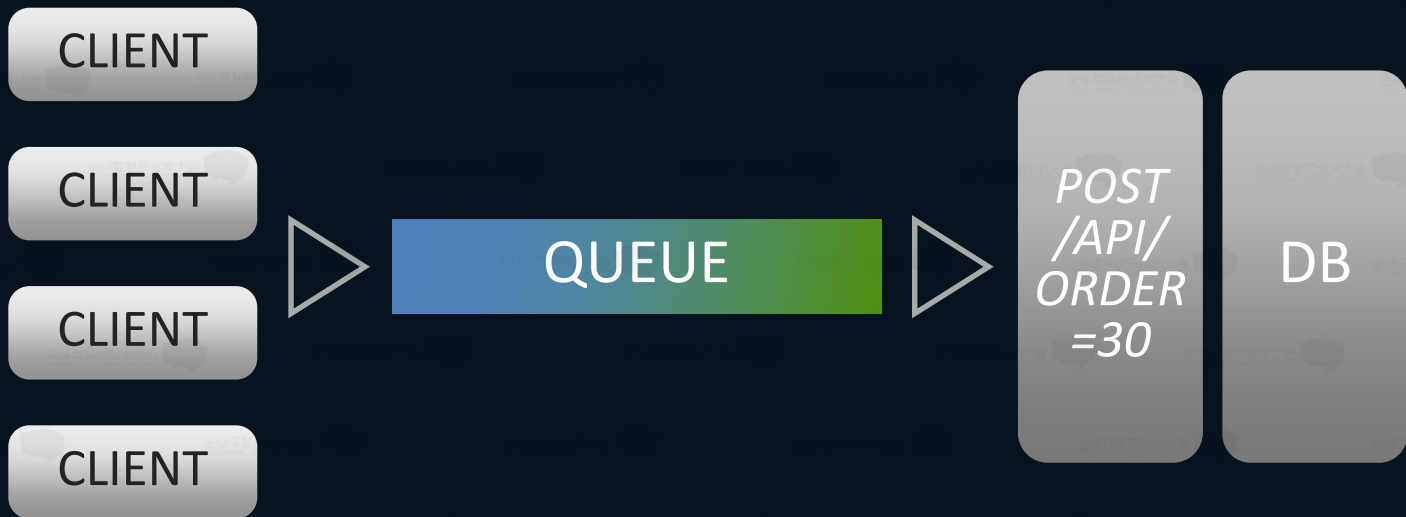
# 数据开放的核心技术

## ➤ 防范攻击类型

负载消耗攻击	HTTP flood
CPU消耗攻击	随机构造攻击
	404穿透攻击
连接数消耗攻击	HTTP发包慢速攻击
	HTTP收包慢速攻击
流量消耗攻击	异常header包攻击
	异常body包攻击
其他	高频验证码校验
	UA保护

# 数据开放的核心技术

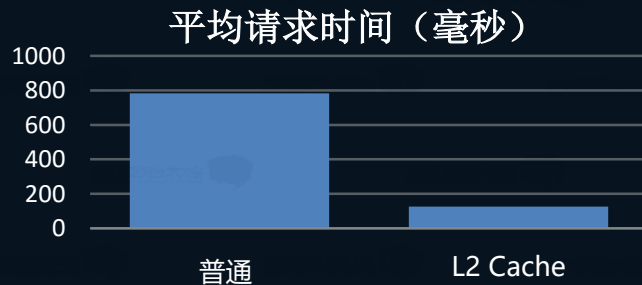
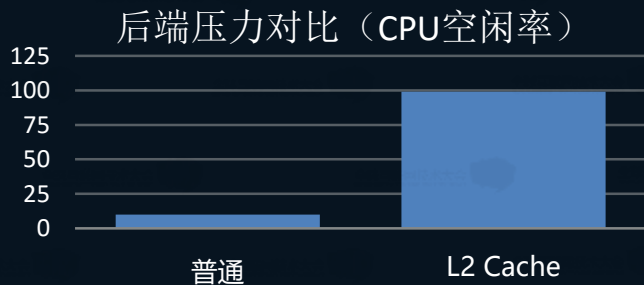
- API加速
  - L2 Cache
  - 接口异步化



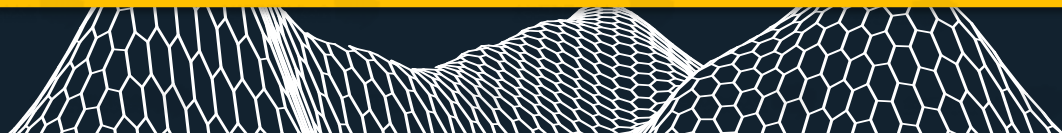
# 数据开放的核心技术

## ➤ API加速效果

GET [HTTP://www.bsclink.com/stock.php?uid=987362&code=300104](http://www.bsclink.com/stock.php?uid=987362&code=300104)



# 4 云聚合产品功能列表



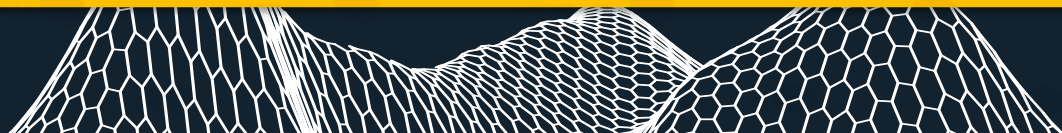
# CLN-X 云聚合

## 接口加速/安全部分 ( 本PPT )

接口加速	L2 Cache
	WebSocket事件驱动转换
接口保护	异步化
	调用者权限
	HTTPS转换
	接口流控
	应用接口防火墙

## 接口治理/管理部分

API构建	API自动生成
	swagger接口定义
	文档生成
	SDK生成
API网关	API网关
	域名管理
	API适配转换
API监控/分析	数据分析
	日志中心
	监控报警



# 让数据流动起来 三层S03

